Frailty models using R

Ørnulf Borgan
Department of Mathematics
University of Oslo

NORBIS course
University of Oslo
4-8 December 2017

The R library parfm may be used to fit shared frailty models with parametric baseline hazard. Default is a Weibull baseline parameterized as $\alpha(t) = \lambda \rho t^{\rho-1}$.

When using a gamma frailty, it is assumed to have mean 1 and variance $\delta$.

See the R documentation for other baseline hazards and frailty distributions.

Fitting parametric frailty models for clustered data

Example: Litter-matched rats

The data consist of 50 litters of female rats with 3 rats in each litter. One rat in each litter received a potentially tumorigenic treatment, the other two were controls. The time (in weeks) until tumor occurrence was observed for each rat. Censoring was due to death without tumor or end of study (at 104 weeks when the rats still alive were sacrificed).

Data for 6 of the 50 litters (T is occurrence of tumor, D is death or sacrifice, i.e. censoring)

<table>
<thead>
<tr>
<th>Litter no.</th>
<th>Drug-treated</th>
<th>Control 1</th>
<th>Control 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>101-D</td>
<td>49-T</td>
<td>104-D</td>
</tr>
<tr>
<td>03</td>
<td>104-D</td>
<td>102-D</td>
<td>104-D</td>
</tr>
<tr>
<td>05</td>
<td>104-D</td>
<td>104-D</td>
<td>104-D</td>
</tr>
<tr>
<td>07</td>
<td>77-D</td>
<td>97-D</td>
<td>79-D</td>
</tr>
<tr>
<td>09</td>
<td>89-D</td>
<td>104-D</td>
<td>104-D</td>
</tr>
<tr>
<td>11</td>
<td>88-T</td>
<td>96-T</td>
<td>104-D</td>
</tr>
</tbody>
</table>

Each litter is a cluster, and the rats are the units. We fit a gamma frailty model (where the frailties have mean 1 and variance $\delta$) with treatment as covariate, assuming a Weibull baseline hazard:

$$\alpha_{ij}(t \mid Z_i) = Z_i \cdot \alpha_{ij}(t) = Z_i \cdot \lambda \rho t^{\rho-1} \exp(\beta x_{ij})$$

Here $x_{ij} = 1$ if rat $j$ in litter $i$ is treated, $x_{ij} = 0$ otherwise.

ML-estimates (with standard errors when year is time unit):

$$\hat{\rho} = 3.93 \ (0.56) \quad \hat{\lambda} = 0.020 \ (0.008)$$
$$\hat{\beta} = 0.908 \ (0.32) \quad \hat{\delta} = 0.489 \ (0.469)$$

Note that the variance is denoted «theta» in the output from parfm.
We will use the likelihood ratio test, to test the null hypothesis that there is no frailty.

This corresponds to testing if the frailty variance is 0.

The usual properties for the likelihood ratio test do not apply in this situation, since the null hypothesis is at the boundary of the parameter space.

One may show that in such situations two times the difference in log-likelihoods is approximately distributed as 
\[ \frac{1}{2} \chi^2_0 + \frac{1}{2} \chi^2_1 \]

To get the correct P-value, we should therefore simply halve the P-value we obtain from the usual likelihood ratio test.

For the rat data, the log-likelihood for the model with frailty is -83.423.

The model without frailty has log-likelihood -84.277.

The likelihood ratio statistic takes the value 
\[ 2(-83.423 + 84.227) = 1.608 \]

The corrected P-value becomes 
\[ 0.2048/2 = 0.102 \]

Thus there is not a significant litter effect for the rat data.
A semi-parametric Cox frailty model

When the baseline hazard in parametric, we have a fully parametric model and the usual likelihood results apply.

We may alternatively consider a Cox type frailty model with a non-parametric baseline:

\[ \alpha_j(t \mid Z_i) = Z_i \cdot \alpha_0(t) \exp(\beta^T x_{ij}) \]

One then considers an extended model where the cumulative baseline hazard \( A_i(t) \) has jumps at the observed event times (only) and maximizes the likelihood with respect to the regression coefficients, the parameters of the frailty distribution and the jumps of the cumulative baseline hazard.

The theory for such semi-parametric models becomes quite involved, but it has been worked out.

Using R

The results of the previous slide are obtained by the commands below.

```r
# Fit a Cox frailty model with gamma frailty (with mean 1 and variance theta) # with treatment (rx) is a covariate::
fit.cox=coxph(Surv(time/52,status)~factor(rx)+frailty(litter), data=female.rats)
summary(fit.cox)

# We obtain the output (edited):

<table>
<thead>
<tr>
<th></th>
<th>coef</th>
<th>se(coef)</th>
<th>se2</th>
<th>Chisq</th>
<th>DF</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor(rx)1</td>
<td>0.914</td>
<td>0.323</td>
<td>0.319</td>
<td>8.01</td>
<td>1.0</td>
<td>0.0046</td>
</tr>
<tr>
<td>frailty(litter)</td>
<td>17.69</td>
<td>14.4</td>
<td>0.2400</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exp(coef)</td>
<td>2.5</td>
<td>0.401</td>
<td>1.32</td>
<td>4.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variance of random effect = 0.499  I-likelihood = -180.8
Degrees of freedom for terms = 1.0  14.4
Concordance = 0.791  (se = 0.048 )
Likelihood ratio test= 37.6 on 15.4 df,  p=0.00124

When interpreting the output we will only consider the numbers given in red
(To test if there is a significant frailty effect, one may compare the "I-likelihood" with the partial likelihood from a Cox model without frailty.)

For the rat data we consider the model

\[ \alpha_j(t \mid Z_i) = Z_i \cdot \alpha_0(t) \exp(\beta x_{ij}) \]

Here \( x_{ij} = 1 \) if rat \( j \) in litter \( i \) is treated, \( x_{ij} = 0 \) otherwise.

The estimates for a Cox frailty model are quite similar to those obtained assuming a Weibull baseline:

\[ \hat{\beta} = 0.914 (0.323) \quad \hat{\delta} = 0.499 \]

The Cox frailty model with gamma frailty is fitted by exploiting the equivalence of a frailty model and a penalized partial likelihood approach (see the book by Therneau & Grambsch, Springer, 2000).

This makes parts of the output difficult to appreciate, so we will only consider the parts of the output on the next page that is given in red.